A REMARK ON THE CONVEXITY AND POSITIVE DEFINITENESS CONCERNING HERMITIAN MATRICES

YIK-HOI AU-YEUNG and YIU-TUNG POON

1.   INTRODUCTION

We denote by F the field R of real numbers or the field C of complex numbers, and by $M_n(F)$ the set of all $n \times n$ matrices with elements in F . A matrix $A \in M_n(F)$ is called hermitian if $A = A^*$, where $*$ denotes the conjugate transpose. We denote by $H_n(F)$ the real linear space of all $n \times n$ hermitian matrices with elements in F . A matrix $A \in H_n(F)$ is said to be positive definite, written $A > 0$, (respectively positive semidefinite, written $A \geqq 0$) if $xAx^* > 0$ (respectively $xAx^* \geqq 0$) for all $x \in F^n \backslash \{0\}$. The purpose of this note is to consider the relationship between the convexity and definiteness concerning hermitian matrices and some related topics. In the following content, we usually assume $n \geqq 2$ and for the convexity argument we identify the complex plane with R^2 .

2.   CONVEXITY AND DEFINITENESS

Let $A \in M_n(C)$, Hausdorff [11] proved that the numerical range of A defined by

$$W(A) = \{xAx^* : x \in C^n , \|x\| = 1\}$$

is convex. If we write $A = A_1 + iA_2$, the unique hermitian decomposition of A , then

$$W(A) = W(A_1, A_2) = \{(xA_1x^*, xA_2x^*) : x \in C^n \text{ and } xx^* = 1\} .$$

Therefore, Hausdorff's result may be stated as:

For any $A_1, A_2 \in H_n(C)$, $W(A_1, A_2)$ is convex.

There were many other proofs of this result, for example see Donoghue [6], Goldman and Marcus [10], Stone [16]. Brickman [4] considered the real analog of the above result. He proved that if $n \geqq 3$, then for any $S_1, S_2 \in H_n(R)$, the set $\{(xS_1x^*, xS_2x^*): x \in R^n \text{ and } xx^* = 1\}$ is convex. (Here $*$ merely means transpose.) Combining these two results, we have

THEOREM 1. *If* $F = C$ *and* $n \geq 2$ *or* $F = R$ *and* $n > 2$, *then, for any* A_1, $A_2 \in H_n(F)$, *the set*

$$W(A_1, A_2) = \{(xA_1x^*, xA_2x^*): x \in F^n \text{ and } xx^* = 1\}$$

is convex.

A unified proof of Theorem 1 was given by Au-Yeung [1].

Let A_1, $A_2 \in M_n(F)$. Then another problem aries: under what condition we can conclude that there exist real numbers α_1 and α_2 such that $\alpha_1 A_1 + \alpha_2 A_2$ is positive definite. To this problem, we have the following theorem.

THEOREM 1'. *If* $F = C$ *and* $n \geq 2$ *or* $F = R$ *and* $n > 2$, *then, for any* A_1, $A_2 \in H_n(F)$ *such that*

$$(xA_1x^*, xA_2x^*) \neq (0, 0) \quad \text{for all } x \in F^n \backslash \{0\} \quad,$$

there exist α_1, $\alpha_2 \in R$ *such that*

$$\alpha_1 A_1 + \alpha_2 A_2 > 0 .$$

Theorem 1' was first proved by Finsler [7] for $F = R$ and there are many other proofs, for example, Calabi [5], Hestenes [12], Taussky [17]. Au-Yeung [2] has given a unified proof of Theorem 1' for both cases of F. For a more detailed survey of this problem, see Uhlig [19].

In [17], Taussky proved that Theorem 1 implies Theorem 1'. But, it seems to the authors that nothing has been mentioned about the converse, which, as we shall see, is also true in a more general form.

A natural generalization of Theorem 1 is to consider, for three A_1, A_2 and $A_3 \in H_n(F)$, the convexity of the set

$$W(A_1, A_2, A_3) = \{(xA_1x^*, xA_2x^*, xA_3x^*): x \in F^n \text{ and } xx^* = 1\} .$$

For $F = C$, Hausdorff [11] mentioned that we can use Toeplitz's method [18] to show that $W(A_1, A_2, A_3)$ has a convex boundary. He also remarked that, in general, $W(A_1, A_2, A_3)$ is not convex. In [4], Brickman gives an example of A_1, A_2, and $A_3 \in H_n(R)$ such that $W(A_1, A_2, A_3)$ is not convex. However, we shall see that for $F = C$ and $n > 2$, $W(A_1, A_2, A_3)$ is always convex.

Let $f_F(n)$ be the dimension of the real linear space $H_n(F)$. Then obviously,

$$f_F(n) = \begin{cases} \dfrac{n(n+1)}{2} & \text{if } F = R \\[2ex] n^2 & \text{if } F = C \ . \end{cases}$$

Now, we can give a generalization of Theorem 1.

THEOREM 2. *If* $1 \leq r \leq n-1$ *and* $p < f_F(r+1) - \delta_{n,r+1}$, *then, for any* $A_1, \ldots, A_p \in H_n(F)$, *the set*

$$W^{(r)}(A_1,\ldots,A_p) = \{(\sum_{i=1}^{r} x_i A_1 x_i^*, \ldots, \sum_{i=1}^{r} x_i A_p x_i^*) : x_i \in F^n \text{ and } \sum_{i=1}^{r} x_i x_i^* = 1\}$$

is convex, where δ_{ij} *is the Kronecker delta.*

We shall proved that Theorem 2 is equivalent to Bohnenblust's result [3], (see Friedland and Loewy [9] for another proof), which is a generalization of Theorem 1'.

THEOREM 2'. (Bohnenblust) *If* $1 \leq r \leq n-1$ *and* $p < f_F(r+1) - \delta_{n,r+1}$, *then, for any* $A_1, \ldots, A_p \in H_n(F)$ *such that*

$$(*) \qquad (\sum_{i=1}^{r} x_i A_1 x_i^*, \ldots, \sum_{i=1}^{r} x_i A_p x_i^*) \neq (0, \ldots, 0)$$

$$\textit{for all} \quad (x_1, \ldots, x_r) \in \underbrace{F^n \times \cdots \times F^n}_{r\text{-times}} \backslash \{(0, \ldots, 0)\}$$

then there exist $\alpha_1, \ldots, \alpha_p \in R$ *such that*

$$\sum_{j=1}^{p} \alpha_j A_j > 0 \ .$$

PROOF.

"Theorem 2 $\Rightarrow$ Theorem 2'"

Since $W^{(r)}(A_1, \ldots, A_p)$ is convex and from condition $(*)$, $W^{(r)}(A_1, \ldots, A_p)$ does not contain the origin in R^p. The result follows from using the Separation Theorem for convex set (for example see [14]).

"Theorem 2' $\Rightarrow$ Theorem 2"

Let $D = \{(x_1, \ldots, x_r): x_i \in F^n$ for $i = i,\ldots,r$ and $\sum_{i=1}^{r} x_i x_i^* = 1\}$.

Define $g: D \rightarrow R^p$ by $g(x_1, \ldots, x_r) = (\sum_{i=1}^{r} x_i A_1 x_i^*, \ldots, \sum_{i=1}^{r} x_i A_p x_i^*)$. Then

$g(D) = W^{(r)}(A_1, \ldots, A_p)$ and we are going to prove that $g(D)$ is convex.

Assume the contrary that $g(D)$ is not convex. Let $\underline{x} = (x_1, \ldots, x_r)$, $\underline{y} = (y_1, \ldots, y_r) \in D$ and $0 < \delta < 1$ such that

$$g(\underline{x}) = (a_1, \ldots, a_p)$$
$$g(\underline{y}) = (b_1, \ldots, b_p)$$

and

$$\delta g(\underline{x}) + (1 - \delta)g(\underline{y}) \notin g(D) .$$

Let I_n be the $n \times n$ identity matrix. Putting

$$B_j = A_j - [\delta a_j + (1-\delta)b_j]I_n \qquad \text{for} \quad j = 1, \ldots, p$$

then $(\sum_{i=1}^{r} z_i B_1 z_i^*, \ldots, \sum_{i=1}^{r} z_i B_p z_i^*) \neq (0, 0, \ldots, 0)$ for all $(z_1, \ldots, z_r) \neq (0, \ldots, 0)$. Therefore by Theorem 2', there exist $\alpha_1, \ldots, \alpha_p \in R$ such that

$$B = \sum_{j=1}^{p} \alpha_j B_j > 0 .$$

But

$$(\sum_{i=1}^{r} x_i B_1 x_i^*, \ldots, \sum_{i=1}^{r} x_i B_p x_i^*) = (1 - \delta)(a_1 - b_1, \ldots, a_p - b_p)$$

and

$$(\sum_{i=1}^{r} y_i B_1 y_i^*, \ldots, \sum_{i=1}^{r} y_i B_p y_i^*) = -\delta(a_1 - b_1, \ldots, a_p - b_p) .$$

Taking the inner product with $(\alpha_1, \ldots, \alpha_p)$ we have

$$(1-\delta)\sum_{j=1}^{p}\alpha_j(a_j - b_j) = \sum_{i=1}^{r} x_i B x_i^* > 0$$

and

$$-\delta \sum_{j=1}^{p} \alpha_j(a_j - b_j) = \sum_{i=1}^{r} y_i B y_i^* > 0$$

which is a contradiction.

The following corollary follows immediately from Theorem 2.

COROLLARY 1. *Let* $n > 2$. *Then for any* $A_1, A_2, A_3 \in H_n(C)$, *the set*

$$\{(xA_1x^*, xA_2x^*, xA_3x^*): x \in C^n \text{ and } xx^* = 1\}$$

is convex.

From Corollary 1, we have

COROLLARY 2. *Let* $n > 2$. *Then for any* $A_1, A_2, A_3 \in H_n(C)$ *and* $\delta \in R$, *the sets*

$$W_1(A_1, A_2; A_3, \delta) = \{(xA_1x^*, xA_2x^*): x \in C^n \text{ and } xx^* = 1 \text{ and } xA_3x^* \geq \delta\}$$

$$W_2(A_1, A_2; A_3, \delta) = \{(xA_1x^*, xA_2x^*): x \in C^n \text{ and } xx^* = 1 \text{ and } xA_3x^* > \delta\}$$

$$W_3(A_1, A_2; A_3, \delta) = \{(xA_1x^*, xA_2x^*): x \in C^n \text{ and } xx^* = 1 \text{ and } xA_3x^* = \delta\}$$

are convex.

By using the hermitian decomposition and the fact that $\|xA\|^2 = xAA^*x^*$, from Corollary 2, we have

COROLLARY 3. *Let* $n > 2$. *Then for any* T , $A \in M_n(C)$ *and any* $\delta \in R$, *the set*

$$V(T; A, \delta) = \{xTx^*: x \in C^n \text{ and } \|x\| = 1 \text{ and } \|xA\| \geq \delta\}$$

is convex.

The convexity of $V(T; A, \delta)$ was first asked by Stampfli [15] for the case $A = T$, and it is recently solved by Kyle [13] for general A and even for the case $n = 2$.

3. RELATED TOPICS

Brickman [4] has proved that for any $A_1, A_2, A_3 \in H_n(C)$, the set

$$\hat{W}^{(1)}(A_1, A_2, A_3) = \{(xA_1x^*, xA_2x^*, xA_3x^*): x \in C^n\}$$

is convex. Recently, this has also been obtained by Fox [8]. It is not difficult to see that this result follows immediately from Corollary 1, except for $n = 2$. We shall generalize Brickman's result. For any $A_1, \ldots, A_p \in H_n(F)$ and any $1 \leq r \leq n-1$, let

$$\hat{W}^{(r)}(A_1, \ldots, A_p) = \{(\sum_{i=1}^{r} x_i A_1 x_i^*, \ldots, \sum_{i=1}^{r} x_i A_p x_i^*); \ x_1, \ldots, x_r \in F^n\}$$

For fixed pair of r, p and $A_1, \ldots, A_p \in H_n(F)$ the convexity of $W^{(r)}(A_1, \ldots, A_p)$ would imply the convexity of $\hat{W}^{(r)}(A_1, \ldots, A_p)$. Given $1 \leq r \leq n-1$, a natural question is to find the maximum number p such that $\hat{W}^{(r)}(A_1, \ldots, A_p)$ is convex for any $A_1, \ldots, A_p \in H_n(F)$. The following theorem shows that even though $\hat{W}$ has the relaxation of $\sum_{i=1}^{r} x_i x_i^* = 1$, the convexity of W and $\hat{W}$ only differs in the case when $r = n-1$.

THEOREM 3. *If* $1 \leq r \leq n-1$ *and* $p < f_F(r+1)$, *then for any* $A_1, \ldots, A_p \in H_n(F)$, *the set* $\hat{W}^{(r)}(A_1, \ldots, A_p)$ *is convex.*

PROOF. Let $1 \leq r \leq n-1$, $p \leq f_F(r+1) - 1$ and $A_1, \ldots, A_p \in H_n(F)$. Since a matrix $X \in H_n(F)$ is ≥ 0 and of rank $\leq r$ if any only if there are $x_1, \ldots, x_r \in F^n$ such that $X = \sum_{i=1}^{r} x_i^* x_i$, we see that

$$\hat{W}^{(r)}(A_1, \ldots, A_p) = \{(Tr\, A_1 X, \ldots, Tr\, A_p X): X \geq 0 \text{ and } rank\, X \leq r\}$$

where 'Tr' means trace.

In order to prove that $\hat{W}^{(r)}(A_1, \ldots, A_p)$ is convex, it suffices to show that for any $X \geq 0$ and $rank\, X = \ell + 1$, where $\ell \geq r$, there exists $Y \geq 0$ and $rank\, Y \leq \ell$ such that

$$Tr\, A_i X = Tr\, A_i Y \qquad \text{for} \quad i = 1, \ldots, p \quad.$$

Without loss of generality, we may assume

$$X = \left[\begin{array}{c|c} I_{\ell+1} & 0 \\ \hline 0 & 0 \end{array} \right]$$

Let

$$\mathcal{L} = \left\{ \left[\begin{array}{c|c} B & 0 \\ \hline 0 & 0 \end{array} \right] : B \in H_{\ell+1}(F) \right\}$$

Then $\dim \mathcal{L} = f_F(\ell + 1) \geq f_F(r + 1)$.

90

Let

$$\mathcal{L}(A_1, \ldots, A_p)^\perp = \{X \in H_n(F) : \operatorname{Tr} A_i X = 0 \ \text{ for } \ i = 1, \ldots, p\} \quad .$$

Then $\dim \mathcal{L}(A_1, \ldots, A_p)^\perp \geqq f_F(n) - p \geqq f_F(n) - f_F(r+1) + 1$. Hence

$$\mathcal{L} \cap \mathcal{L}(A_1, \ldots, A_p)^\perp \neq 0 \quad .$$

Let $\tilde{B} = \begin{bmatrix} B & 0 \\ \hline 0 & 0 \end{bmatrix} \in \mathcal{L} \cap \mathcal{L}(A_1, \ldots, A_p)^\perp \setminus \{0\}$. We may assume the eigenvalues of

B are $1 = \beta_1 \geqq \cdots \geqq \beta_{\ell+1}$.

So $Y = X - \tilde{B} \geqq 0$, rank $Y \leqq \ell$ and

$$\operatorname{Tr} A_i Y = \operatorname{Tr} A_i X \quad \text{for all} \quad i = 1, \ldots, p \ .$$

This completes the proof of Theorem 3.

REMARKS.

1. The bounds $f_F(r+1) - \delta_{n,r+1}$ and $f_F(r+1)$ in Theorems 2, 2' and 3 respectively are best possible.

2. If we consider the skew field Q of quaternions, then Theorems 2, 2' and 3 remain valid if we take $f_Q(n) = 2n^2 - n$.

REFERENCES

1. Y.H. Au-Yeung, A simple proof of the convexity of the field of values defined by two hermitian forms, Aequationes Math. 12, 82–83 (1975).

2. Y.H. Au-Yeung, A theorem on a mapping from a sphere to the circle and the simultaneous diagonalization of two hermitian matrices, Proc. Amer. Math. Soc. 20, 545–548 (1969).

3. F. Bohnenblust, Joint positiveness of matrices, unpublished manuscript.

4. L. Brickman, On the field of values of a matrix, Proc. Amer. Math. Soc. 12, 61–66 (1961).

5. E. Calabi, Linear systems of real quadratic forms, Proc. Amer. Math. Soc. 15, 844–846 (1964).

6. W.F. Donoghue Jr., On the numerical range of a bounded operator, Michigan Math. J. 4, 261–263 (1957).

7. P. Finsler, Über das Vorkommen definiter und semidefiniter Formen in Scharen quadratischer Formen, Comment. Math. Helv. 9, 188–192 (1936/37).

8. D.W. Fox, The convexity of the range of three hermitian forms and of the numerical range of sesquilinear forms, Linear Algebra Appl. 22, 191–194 (1978).

9. S. Friedland and R. Loewey, Subspaces of symmetric matrices containing matrices with a multiple first eigenvalue, Pacific J. Math. 62, 389–399 (1976).

10. A.J. Goldman and M. Marcus, Convexity of the field of a linear transformation, Canad. Math. Bull. 2, 15–18 (1959).

11. F. Hausdorff, Der Wertvorrat einer Bilinearform, Math. Z. 3, 314–316, (1919).

12. M.R. Hestenes, Pairs of quadratic forms, Linear Algebra Appl. 1, 397–407 (1968).

13. J. Kyle, $W_\delta(T)$ is convex, Pacific J. Math. 72, 483–485 (1977).

14. R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, N.J. 1970, 95–101.

15. J.G. Stampfli, The norm of a derivation, Pacific J. Math. 33, 737–747 (1970).

16. M.H. Stone, Hausdorff's theorem concerning hermitian forms, Bull. Amer. Math. Soc. 36, 259–261 (1930).

17. O. Taussky, Positive-definite matrices, in Inequalities (O. Shisha, Ed.), Academic, New York, 1967, 309–319.

18. O. Toeplitz, Das algebraishe Analogon zu einem Satze von Fejér, Math. Z. 2 (1918), 187–197.

19. F. Uhlig, A recurring theorem about pairs of quadratic forms and extensions: A survey, Linear Algebra Appl. 25, 219–237 (1979).

Department of Mathematics
University of Hong Kong
Hong Kong